

Indirect Evaluation of Text Clusters

- Question to answer: how useful are the clustering results for the intended applications?
 - “Usefulness” is inevitably application specific
 - “Clustering bias” is imposed by the intended application
- Evaluation procedure:
 - Create a test set for the intended application to quantify the performance of any system for this application
 - Choose a baseline system to compare with
 - Add a clustering algorithm to the baseline system → “clustering system”
 - Compare the performance of the clustering system and the baseline in terms of any performance measure for the application